Manuscript prepared for Nonlin. Processes Geophys.
with version 2014/09/16 7.15 Copernicus papers of the LaTeX class copernicus.cls.
Date: 6 May 2015

Spectral Diagonal Ensemble Kalman Filters Response to referees' comments

Ivan Kasanický[1], Jan Mandel[1,2], and Martin Vejmelka[1]

[1]Institute of Computer Science, Academy of Sciences of the Czech Republic
[2]Department of Mathematical and Statistical Sciences, University of Colorado Denver

Correspondence to: Jan Mandel (jan.mandel@gmail.com)

Abstract. We provide responses to both referees.

1 Introduction

We would like to thank the anonymous referees for their comments.

2 Anonymous referee 1

The anonymous referee 1 has 2 minor concerns and 2 minor comments, called "minor stuff".

2.1 Minor concerns

Minor concern 1

Referee's comment:

This manuscript is unusually clear in its presentation of methodology. The method described is shown to be more accurate for certain types of applications. My only concern was that the particular experiments used to demonstrate the method's capabilities seemed to be for unusual parameter regimes, and I would like to understand better why the authors chose these cases. I also note a few minor grammatical errors. I think that the manuscript is acceptable pending these minor amendments. Minor concerns: 1. The authors begin their results section with Lorenz96. This non-dimensional model requires an ad hoc mapping to a dimensional time and there is a standard in the literature for doing this. Normally, experiments assimilate at frequencies that are roughly similar to those in numerical weather prediction when using the standard time definition. For instance, assimilating once an hour would be deemed to be high frequency while every 6 or 12 hours would be common. The authors here in section 7.1 state that they are assimilating once every second after assigning 0.01s for the time step. It is important that the authors clearly describe how they are defining dimensional time and highly preferable that they use the standard definition. It still seems to me that they are assimilating very frequently compared to most applications in the literature, and if so, this should be motivated. Similarly, the observational error variance of 0.01 is very small compared to most published applications for L96. This error variance is a tiny fraction of the "climatological" variance of L96, and doesn't look like a very reasonable analogy for the types of error variances found in real geophysical applications. Again, the authors should clearly state why they used such a small value and comment on the relation to the more common values in the geophysical literature.

Similarly, the assimilation frequency for the shallow water example seemed odd. Assimilating every second is very frequent compared to the evolution of the dynamics. Some readers may become suspicious that high frequency assimilation was chosen as a case for which the new method is particularly competitive. Again, the authors should clearly state why they picked such frequent assimilation and how they picked the observational error variance.

Author's response:

Our description of an experiment with the Lorenz 96 model in the discussion paper incorrectly assumed that the time unit was the second. We have assimilated into the Lorenz 96 model not every second, but every time unit. According to Lorenz and Emanuel (1998), 0.05 time unit of Lorenz model is equivalent to 6 hours of climatological model. Since we assimilated every 1 time unit, this is equivalent to assimilation every 5 days.

To test our proposed method with standard parameterization we used parameter values introduced in Lorenz and Emanuel (1998), in particular their observation error variance, with only the small difference that the length of state vector was 256 and we used time step 0.01 for model evaluation.

In the experiments with shallow water equation model, we extended the assimilation cycle to 1 hour, compared to 1 minute used in the discussions manuscript. Since we used a very simple model, the model and numerical methods used to solve start to degenerate after 15-20 hours.

The changes in experiment setups have no significant impact on results and our proposed method still works much better then standard EnKF.

Changes in manuscript:

We rewrote sections 7.1 and 7.2, where the experiments are described, according to the new design of experiments. We also prepared new figures with correct labels for the manuscript.

Minor concern 2:

Referee's comment:

Not required, but would be a nice addition: There is limited discussion of how the method extends to non-identity forward operators and none about nonlinear forward operators. A paragraph in the conclusions would be a nice addition if something simple can be said.

Author's response:

The method described in Section 6.3 allows general linear operators, but inverse (i.e., solving a system) in the observation space is required. This issue is well known in spectral variational methods; techniques used in the literature include aggregating and interpolating observations to create "super-obs" gridded arrays (Parrish and Derber, 1992).

Changes in manuscript:

We will add a paragraph like the above to the conclusion. Nonlinear observation operators remain for future work.

2.2 Other comments

Minor stuff 1:

Referee's comment:

p. 6, line 11: $N \ll n$. Real issue is whether $N \ll q$, where q is the rank of the covariance matrix of the Kalman filter solution. This is an important issue because this mistake has repeatedly confused things in the geophysical ensemble literature.

Author's response:

This is a more or less standard introductory statement, not original research in any way. The rank of the state covariance (which is the covariance from the KF solution) is not smaller than the dimension unless the problem is degenerate, which is the focus neither in this paper nor in applications. Rather, there are usually many small but positive eigenvalues. The reviewer probably means by q the number of significant modes, i.e., the number of eigenvalues above some positive threshold, i.e., effective rank.

Changes in manuscript:

We will replace "$N \ll n$" by something like "N much less than the number of significant modes" but we will not explain what modes are. We will add a standard comment that the low rank causes spurious long-long range correlations.

Minor stuff no. 2:

Referee's comment:

p. 6, line 15: 'transform' to 'transforms'

Author's response:

Changes in manuscript:

Replaced.

Minor stuff no. 3:

Referee's comment:

. p. 7, line 16: 'Eq. 5' should be 'Eq. 7'

Author's response:

Changes in manuscript:

Replaced.

Minor stuff no. 4:

Referee's comment:

p. 8, line 20: It would be better to put the equation that starts at the end of this page on its own line.

Author's response:

Page breaks will change, the final version has regular-sized pages.

Changes in manuscript:

We will put the equation on a displayed line of its own so that it does not break between lines (and possibly pages).

Minor stuff no. 5:

Referee's comment:

p. 10, line 14: mean IS known?

Author's response:

Changes in manuscript:

Added 'is'.

Minor stuff no. 6:

Referee's comment:

p. 11, line 3: It might be clearer to say 'only one gridded variable'. I've had previous experience with saying 'one variable' and having readers interpret that as a scalar.

Author's response:

Changes in manuscript:

We have replaced the name of the subsection as the referee suggest. We have also replaced "one variable" by "one gridded variable" on line 4.

Minor stuff no. 7:

Referee's comment:

P. 12, line 5: 'is THE one by'

Author's response:

Changes in manuscript:

Added 'the'.

Minor stuff no. 8:

Referee's comment:

P. 12, line 8: 'is THE matrix'

Author’s response:

Changes in manuscript:

Added ‘the’.

Minor stuff no. 9:

Referee’s comment:

P. 12, line 13: ‘by A call to’

Author’s response:

Changes in manuscript:

Added ‘a’.

Minor stuff no. 10:

Referee’s comment:

P. 13, line 5: ‘one for each data point’

Author’s response:

Changes in manuscript:

Replaced ‘points’ by ’point’.

Minor stuff no. 11:

Referee’s comment:

P. 13, line 15: Unclear to me why these must be contiguous. Couldn’t this work for any subset of variables? If not, you might add a sentence to make it clear why (not even clear what ‘continguous’ means for a grid).

Author’s response:

The method goes through for any observed subset of entries of the gridded variable $\mathbf{X}_1$, but the performance will vary. The performance tends to be better when the observed and unobserved entries of $\mathbf{X}_1$ fill two subdomains of the physical domain with a relatively small boundary between them. A detailed investigation, however, is planned for elsewhere.

Changes in manuscript:

We will delete the word ‘contiguous’ and include a brief remark as above.

Minor stuff no. 12:

Referee’s comment:

P. 14, line 16: ‘state consistS’

Author’s response:

Changes in manuscript:

Replaced ‘consist’ by ‘consists’.

Minor stuff no. 13:

Referee’s comment:

P. 14, line 19: ‘minimalizes’ to ‘minimizes’

Author’s response:

Changes in manuscript:

Replaced.

Minor stuff no. 14:

Referee’s comment:

P. 15, line 20: Doesn’t the KF minimize the expected RMSE for linear Gaussian?

Author’s response:

The problem is not linear Gaussian and the paragraph refers to EnKF (which is approximate), not the KF.

Changes in manuscript:

None.

Minor stuff no. 15:

Referee's comment:

P. 15, line 7: 'timestep of THE model'

Author's response:

Changes in manuscript:

Added 'the'.

Minor stuff no. 16:

Referee's comment:

P. 15, line 14: 'but THE spectral'

Author's response:

Changes in manuscript:

Added 'the'.

Minor stuff no. 17:

Referee's comment:

P. 15, line 22: decreaseS the RMS?

Author's response:

Changes in manuscript:

Replaced 'decrease' by 'decreases'.

Minor stuff no. 18:

Referee's comment:

P. 16, line 7: Aren't u and v normally described as velocity components, rather than momentum?

Author's response:

Changes in manuscript:

Corrected.

Minor stuff no. 19:

Referee's comment:

P. 16, line 10: 'where' to 'were'

Author's response:

Changes in manuscript:

Replaced 'where' by 'were' on line 15.

Minor stuff no. 20:

Referee's comment:

P. 17, lines 6-14: Could part of this be coordinated with last paragraph on p. 16?

Author's response:

Changes in manuscript:

The repetition was deleted.

Minor stuff no. 21:

Referee's comment:

21. P. 17, line 23: 'continuous? to 'contiguous'?

Author's response:

Changes in manuscript:

Deleted "continuous".

Minor stuff no. 22:

Referee's comment:

P. 18, line 1: error THAN the sample?

Author's response:

Changes in manuscript:

Replaced 'that' by 'than'.

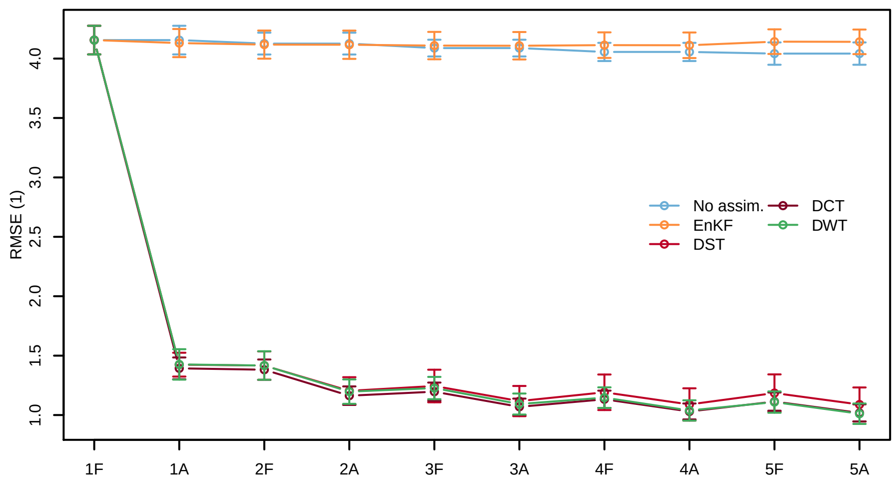

Figure 1. Error plot of 10 repetitions of assimilation of the full state to the Lorenz 96 model, 5 assimilation cycles, F means forecast, A means analysis.

Minor stuff no. 23:

Referee's comment:

Figures: Since you ran multiple realizations, you might want to mention what the error bars would look like (including them seems a bit much).

Author's response:

We include Figs. 1 and 2 with error bars here and in the recvised paper. We can see that the present method performs much better than the standard EnKF.

Changes in manuscript:

3 Anonymous referee 2

Major comment 1

Referee's comment:

The derivation of Theorem 1 (p. 5) is obtained through the spectral decomposition of the matrix, but the present derivation can be obtained directly from the computation in an arbitrary basis since Eq.(10) and Eq.(11) can be rewritten using the intrinsic operator of trace as

$$E[||\mathbf{C}-\mathbf{C}^N||_F^2] = \frac{1}{N-1}\,\mathrm{Trace}(\mathbf{C}^2) + \frac{1}{N-1}\,\mathrm{Trace}(\mathbf{C})^2$$

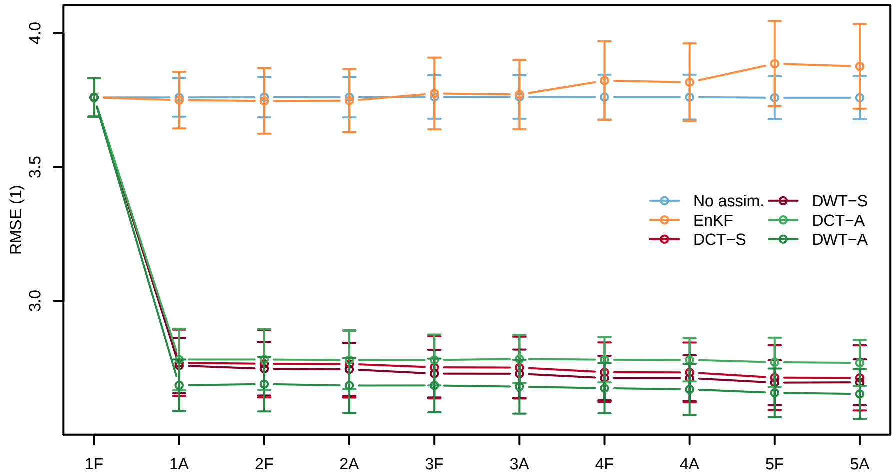

Figure 2. Error plot of 10 repetitions of assimilation of one half of the state vector to the Lorenz 96 model, 5 assimilation cycles, F means forecast, A means analysis.

$$E[||\mathbf{C} - \mathbf{D}^N||_F^2] = \frac{1}{N-1}\mathrm{Trace}(\mathbf{C}^2)$$

independent of the basis and directly related to the spectrum. Hence, the comment p6, l 164 "the analysis in Furrer and Bengtsson (2007) is in the physical domain rather.. " can be suppressed. From my view, this theorem is not really new, and references to Mallat and Furrer & Bengtsson should be enough. If the authors really want to put something here (to make the manuscript self contained), then they should mention the Wick formula that helps to compute general formulation of Gaussian moments. I think enough the derivation when the average is assume known equal to 0.

Author's response:

We would like to thank the referee for the elegant writing of the equalities and the Mallat (1998) reference. We thought of writing the theorem in terms of eigenvalues and without a reference to transformation, and then apply it in the frequency domain where the covariance is diagonal, but then we have eventually decided to streamline the presentation and write it directly for the case we need.

Furrer and Bengtsson (2007) analyze tapering to the diagonal in the physical space, but diagonal covariance in the physical space is never used in applications. The present method is EnKF with diagonal approximation in the frequency domain, where it is reasonable to expect that the covariance will be approximately diagonal. The underlying theorem is similar, but the novelty is in the application. Extending the result in Furrer and Bengtsson (2007, Eqs. 12 and 16) and Mallat (1998, Prop. 10.14 and Eq. 10.179) from zero mean to unknown mean seems also new, and, in applications, the mean is never known. We are not aware of a process to obtain the unknown mean case mathematically from the zero-mean case, even if, in the end, it results in merely replacing N by $N-1$ in the covariance formula, as expected, and we believe that a proof is in order.

Changes in manuscript:

We will add the reference to Mallat (1998, Prop. 10.14 and Eq. 10.179). We will formulate the estimate in general, without reference to a spectral transformation, then apply it in the frequency domain.

Major comment 2

Referee's comment:

As mentioned in the general comments, the method proposed here is not strictly new since in variational data assimilation, algorithms are existing that estimate the covariance of the day, introducing the flow dependence (equivalent to the EnKF) within the cost function minimising process; with quantification of real impact in operational NWP ! Hence, it is important to mention this point in the manuscript (e.g. Buehner, 2005 ; Berre et al., 2007 ; Varella et al., 2011): if the strategy of resolution of the BLUE is different in the hybrid 3DVar and the EnKF, the idea to model the covariance matrix to benefit of a noise less matrix is the same. In the major comment (6) below, I provide you elements to precise this point in the introduction and conclusion of the work. Note also that hybrid formulation comes from EnKF community with the work of Hamill and Snyder (2000) that also have introduced a spectral diagonal assumption (see their Eq.(3)) : this should be specified in the introduction. Of course, with the diagonal assumption in spectral space, since only homogeneous and isotropic correlations can be represented, there is no need to update the diagonal at each analysis step and climatological estimation is better. This is no more true with other formulations as encountered with the wavelets (and frame) that are able to produce heterogeneous correlation function where the spatio-temporal evolution makes sens.

Author's response:

Since major comments 2 to 5 are closely related, we will respond to them together below.

Major comment no. 3

Referee's comment:

The formulation of the background error covariance model using the diagonal assumption following a product of linear operator should mention all the work done in variational literature that intensively relies on this trick to build covariance matrix in huge dimension (Courtier et al. , 1998; Fisher and

Andersson 2001; Weaver and Courtier, 2001). In particular, this should be specified in line 180 where operator transforms (FFT, DWT) are mentioned.

Major comment no. 4

Referee's comment:

From this link with the variational community, some perspectives of the present contribution must be precise. In particular, from the history about covariance modelling in variational algorithm, the next steps of the work can be drawn as follows: construction of non-separable formulation (Courtier et al., 1998; Fisher and Andersson, 2001 ; Pannekoucke, 2009), representation of balances between variables in order to obtain a more realistic multivariate formulation (Derber and Bouttier, 1999 ; Fisher, 2003 ; Weaver et al. 2005), representation of heterogeneity using a physical/spectral localised formulation (non-separable wavelet formulation for Fisher and Anderson, 2001 ; separable formulation based on diffusion operator for Weaver and Courtier, 2001 or recursive filter for Purser et al. 2003 ; nonseparable formulation based on hybridation diffusion/wavelets Pannekoucke, 2009) ... In particular, even if formulations as the diffusion operator or the recursive filter are not diagonal assumptions, they lead to an approximation of the covariance matrix free of sampling noise, and objectively parametrised from ensemble estimation (Pannekoucke and Massart, 2008 ; Michel, 2013 ; Pannekoucke et al. 2014). Along this route, filtering strategies can be employed to improve the estimation and the design of covariance formulations using results on the estimation of variances and length-scales (Berre et al, 2007 ; Raynaud et al. 2009 ; Raynaud and Pannekoucke 2013 ; Menetrier et al. 2015).

Major comment no. 5

Referee's comment:

From the above major comments, saying "The paper provide a new technology for data assimilation" is too much and risks to appear arrogant while considering all the work that has been done for each community. However you right that until now very few person have try to seriously consider covariance model in EnKF, the main reason is that it require to build covariance matrix parameterisation, this represents a real cost in terms of technology investment for NWP codes. You should mention this in the introduction of the paper: "The idea of using covariance model in EnKF to benefit of sample noise reduction effect is known (Hamill and Snyder, 2000; Buehner, 2005), but as far as we know no reference has been published to document the real advantage of this method. In terms of practical implementation of the BLUE, one of the reason could be the relative distance existing between the EnKF and the variational to resolve an equivalent analysis step. However, the employ of forecast ensemble has been tested with success in the variational framework (Buehner 2005, Berre et al. 2007, Varella et al., 2011)". For the conclusion, I guess you can replace the sentence "The paper provide..” as follows: “The paper provide a preliminary test, within an academic setting, of the employ of parametric covariance in pure EnKF strategy, while the reverse strategy is existing in variational framework (hybrid formulation)”.

Author’s response:

Since major comments 2 to 5 above are closely related, we respond to them together.

The novel technique here is the use of spectral covariance modeling in each EnKF analysis cycle to reduce the ensemble size required. The paper Buehner (2005), mentioned by the reviewer, is about improving estimates of background covariance for 3DVAR rather than improving the EnKF. Hamill and Snyder (2000) use a linear combination of sample covariance, different in every analysis cycle, and spectral diagonal covariance from Parrish and Derber (1992), constant in time, rather than spectral modeling in the analysis cycle.

We are aware that spectral diagonal covariance models have been used for estimation of background covariance in variational methods. We have addressed this issue and provided representative references in the discussions paper. We would like to thank the anonymous reviewer very much for providing a much more comprehensive and detailed survey and pointing out important historical milestones, which will contribute greatly to the improvement of the paper. Taking the liberty to paraphrase the material kindly provided by the reviewer, we will write in the introduction and the conclusion something like the following.

Changes in manuscript:

In the introduction: Spectral diagonal covariance models and their estimation from an ensemble of realizations are not new. Spectral diagonal model of background covariance has been used in operational weather forecasting for a long time (Parrish and Derber, 1992). Estimates of background covariance from an ensemble, called flow-dependent covariance, in combinations with spectral covariance models have been used in variational data assimilation (e.g., Buehner, 2005; Buehner and Charron, 2007; Berre et al., 2007; Varella et al., 2011), leading to hybrid EnKF – 3DVAR methods. Another hybrid formulation in EnKF was proposed by Hamill and Snyder (2000), who used a linear combination of sample covariance, different in every analysis cycle, and spectral diagonal covariance from Parrish and Derber (1992), constant in time.

Further developments in the history of background covariance modeling in variational algorithms include construction of non-separable formulation (Courtier et al., 1998; Fisher and Andersson, 2001; Pannekoucke, 2009), representation of balances between variables in order to obtain a more realistic multivariate formulation (Derber and Bouttier, 1999; Fisher, 2003; Weaver et al., 2005), representation of heterogeneity using a physical/spectral localised formulation (non-separable wavelet formulation (Fisher and Andersson, 2001), separable formulation based on diffusion operator (Weaver and Courtier, 2001) or recursive filter for (Purser et al., 2003), and nonseparable formulation based on hybridization of diffusion and wavelets (Pannekoucke, 2009). Formulations such as the diffusion operator or the recursive filter are related to the diagonal assumption here, they involve covariance models with a relatively small number of parameters, thus free of sampling noise, but estimated from an ensemble directly (Pannekoucke and Massart, 2008; Michel, 2013; Pannekoucke et al., 2014). Similar filtering strategies can be employed to improve the estimation and the design of covariance formulations using results on the estimation of variances and length scales (Berre et al., 2007; Raynaud et al., 2009; Raynaud and Pannekoucke, 2013; Ménétrier et al., 2015). The formulation of the background error covariance model using the diagonal assumption and a product of linear operator (such as the discrete Fourier or wavelet transform here) is widely used in variational literature to build covariance models in high dimension (e.g., Courtier et al., 1998; Fisher and Andersson, 2001; Weaver and Courtier, 2001).

The idea of using covariance model in conjunction with EnKF to benefit sample noise reduction is known, but as far as we know no reference has been published to document the real advantage of this method in improvements to the performance of the EnKF. The paper provides a preliminary test, within an academic setting, of the techniques of employing parametric covariance in the EnKF, while the existing literature is focused on the opposite direction, the use of EnKF to provide estimates for the variational framework, known as “hybrid formulation”.

The use of spectral covariance modeling in each EnKF analysis cycle to reduce the ensemble size seems to be new. We are not aware of previous attempts to seriously consider covariance modeling in the EnKF. The main reason could be that it requires to build covariance matrix parameterisation, which represents a real cost in terms of technology investment for NWP codes.

Major comment no. 6

Referee’s comment:

p11, l314 : "because an implementation only needs an orthogonal transformation" orthogonality is not a necessary condition for diagonal assumption that can be considered within a frame as detailed in Pannekoucke et al. (2007). Of course this have an influence for the representation of the observational error covariance matrix in the "spectral space" that is no more diagonal (as specfified in p7, l 179). Note that in Pannekoucke (2008, appendix D) wavelet packets are used to take advantage of the orthogonal basis dictionary it provides ; for this, the problem is then to connect horizontal sheet along the vertical in a 2D/3D formulation, in direct 3D formulation this could be used without the connection issue.

Author’s response:

We state only that the present method can be implemented efficiently using the fast Fourier or wavelet transform. We are aware of the broader issues, such diagonal approximation with frames in Pannekoucke et al. (2007) as well as connecting horizontal sheets with spectral approximation in each. We intend to study these issues in future work.

Changes in manuscript:

We will add at the end on conclusion something like the following: The present method uses orthogonal transformation, but orthogonality is not a necessary condition for diagonal assumption; diagonal approximation with frames was proposed in Pannekoucke et al. (2007). The question of further reducing the number of parameters and thus sampling noise, as in, e.g., functions of the Laplace operator, is also of interest. When spectral diagonalization is used in the horizontal planes, the question is how to connect horizontal sheets along the vertical dimension. In Pannekoucke (2008, Appendix D), wavelet packets are used to take advantage of the orthogonal basis dictionary they provide. These issues will be studied elsewhere.

3.1 Minor comments

Minor comment no. 1

Referee's comment:

p2, l54: Add references for the wavelets formulation: Fisher and Andersson, 2001; Deckmyn and Berre, 2005.

Author's response:

Changes in manuscript:

We will change "The Fourier diagonalization approach was extended by Pannekoucke et al. (2007) to sparse representation of the background covariance by thresholding wavelet coefficients" to "The Fourier diagonalization approach was extended by Pannekoucke et al. (2007) to sparse representation of the background covariance by wavelets (Fisher and Andersson, 2001; Deckmyn and Berre, 2005; Pannekoucke et al., 2007)."

Minor comment no. 2

Referee's comment:

p6, before sec. 6: You should mention that the parametric formulation does not converge toward the "true" covariance matrix as the ensemble size increases toward infinity.

Author's response:

Changes in manuscript:

We will add at the end of Section 5: While the spectral diagonal formulation improves the approximation for small ensembles, it should be noted that it does not converge to the covariance as $N \to \infty$, unless the covariance is diagonal in the spectral basis.

Minor comment no. 3

Referee's comment:

It is not direct that T defined p10, l 283 is positive, please provide a proof of it

Author's response:

Positive definiteness could be verified by counting the eigenvalues of matrix T. Another approach could be by writing T using Kronecker product $T = K \otimes (M \otimes M)$, where K is 3×3 square matrix with elements $(K)_{i,i} = 1$, $(K)_{i,j} = 0.9$ if $i \neq j$ and M is 64×64 square matrix with elements $(M)_{i,j} = \exp(|i-j|)$. All eigenvalues of T are of the form $\alpha.\beta.\gamma$, where α is eigenvalue of K, β and γ are eigenvalues of M. Since the smallest eigenvalues of K and M are 0.1 and 0.41 respectively, all eigenvalues of T are positive.

Changes in manuscript:

We will add: "Note that matrix T could be also written using Kronecker product as $T = K \otimes (M \otimes M)$, where K is 3×3 square matrix with elements $(K)_{i,i} = 1$, $(K)_{i,j} = 0.9$ if $i \neq j$ and M is 64×64 square matrix with elements $(M)_{i,j} = \exp(|i-j|)$. Since both matrices K and M are positive definite, matrix T is also positive definite."

Minor comment no. 4

Referee's comment:

P9, l250: 's' is meaningless here since 't' is a pseudo-time (Lorenz'96 is not related to a physical model but only an academic framework, nice to play with), hence replace "0.01s" by 0.01 and $1s$ by 1

Author's response:

Changes in manuscript:

We will change the design of experiments according to suggestion from anonymous referee 1, therefore we also change Section 7.1 and all figures captions and labels (please see also minor comment from anonymous referee 1).

Minor comment no. 5

Referee's comment:

p9, l 252: "The the ensemble .." → "Then the .."

Author's response:

We did not find the mistake, but we found mistake on p .128 l.22 and replaced "of the the standard" by "of the standard".

Changes in manuscript:

Minor comment no. 6

Referee's comment:

p10, l297 : "To relax the ensemble members the model .. " must be rephrased

Author's response:

Changes in manuscript:

The description in section 7.2 will be rewritten.

Minor comment no. 7

Referee's comment:

p11, l 310 : "experimens" → experiments

Author's response:

Changes in manuscript:

Replaced.

Minor comment no. 8

Referee's comment:

p11, l 311 : "shown that the method that the analysis" must be rephrased

Author's response:

Changes in manuscript:

Deleted "the method".

Minor comment no. 9

Referee's comment:

p13, l 345 : the parenthesis are not well positioned leading to ambiguities, please write something as $E[(\sum \ldots)^2]$, this appears at many times in the derivation of the proof.

Author's response:

Changes in manuscript:

We changed the notation as the referee suggest, all the terms $E[\ldots]^2$ in the appendix were replaced by $E[(\ldots)^2]$. We also changed the notation on line 141 of discussions manuscript, where it could lead to the same ambiguity.

Minor comment no. 10

Referee's comment:

Some work exists concerning the balance in EnKF that should be mention in the manuscript see Kepert (2009).

Author's response:

Changes in manuscript:

We have added in the introduction "Balanced update and localization in the EnKF using the stream function-velocity potential representation were studied in Kepert (2009).".

4 Conclusions

We have addressed all comments by the anonymous reviewers.

References

Berre, L., Pannekoucke, O., Desroziers, G., Stefanescu, S., Chapnik, B., and Raynaud, L.: A variational assimilation ensemble and the spatial filtering of its error covariances: increase of sample size by local spatial averaging, http://old.ecmwf.int/publications/library/ecpublications/_pdf/workshop/2007/Data_assimilation/Berre.pdf, 2007.

Buehner, M.: Ensemble-derived stationary and flow-dependent background-error covariances: Evaluation in a quasi-operational NWP setting, Quarterly Journal of the Royal Meteorological Society, 131, 1013–1043, doi:10.1256/qj.04.15, 2005.

Buehner, M. and Charron, M.: Spectral and spatial localization of background-error correlations for data assimilation, Quarterly Journal of the Royal Meteorological Society, 133, 615–630, doi:10.1002/qj.50, 2007.

Courtier, P., Andersson, E., Heckley, W., Vasiljevic, D., Hamrud, M., Hollingsworth, A., Rabier, F., Fisher, M., and Pailleux, J.: The ECMWF implementation of three-dimensional variational assimilation (3D-Var). I: Formulation, Quarterly Journal of the Royal Meteorological Society, 124, 1783–1807, doi:10.1002/qj.49712455002, 1998.

Deckmyn, A. and Berre, L.: A Wavelet Approach to Representing Background Error Covariances in a Limited-Area Model, Monthly Weather Review, 133, 1279–1294, doi:10.1175/MWR2929.1, 2005.

Derber, J. and Bouttier, F.: A reformulation of the background error covariance in the ECMWF global data assimilation system, Tellus, 51A, 195–221, doi:10.1034/j.1600-0870.1999.t01-2-00003.x, 1999.

Fisher, M.: Background error covariance modelling, in: Proceedings of the ECMWF Seminar on Recent Developments in Data Assimilation for Atmosphere and Ocean, pp. 45–63, ECMWF, Reading, UK, 2003.

Fisher, M. and Andersson, E.: Developments in 4D-Var and Kalman Filtering, ECMWF Technical Memorandum 347, available at http://www.ecmwf.int/publications/library/ecpublications/_pdf/tm/301-400/tm347.pdf, retrieved November 2013, 2001.

Furrer, R. and Bengtsson, T.: Estimation of high-dimensional prior and posterior covariance matrices in Kalman filter variants, J. Multivariate Anal., 98, 227–255, doi:10.1016/j.jmva.2006.08.003, 2007.

Hamill, T. M. and Snyder, C.: A Hybrid Ensemble Kalman Filter–3D Variational Analysis Scheme, Monthly Weather Review, 128, 2905–2919, doi:10.1175/1520-0493(2000)128<2905:AHEKFV>2.0.CO;2, 2000.

Kepert, J. D.: Covariance localisation and balance in an Ensemble Kalman Filter, Quarterly Journal of the Royal Meteorological Society, 135, 1157–1176, doi:10.1002/qj.443, 2009.

Lorenz, E. N. and Emanuel, K. A.: Optimal Sites for Supplementary Weather Observations: Simulation with a Small Model, Journal of the Atmospheric Sciences, 55, 399–414, doi:10.1175/1520-0469(1998)055<0399:OSFSWO>2.0.CO;2, 1998.

Mallat, S.: A wavelet tour of signal processing, Academic Press, Inc., San Diego, CA, 1998.

Ménétrier, B., Montmerle, T., Michel, Y., and Berre, L.: Linear Filtering of Sample Covariances for Ensemble-Based Data Assimilation. Part I: Optimality Criteria and Application to Variance Filtering and Covariance Localization, Monthly Weather Review, doi:10.1175/MWR-D-14-00157.1, 2015.

Michel, Y.: Estimating deformations of random processes for correlation modelling in a limited area model, Quarterly Journal of the Royal Meteorological Society, 139, 534–547, doi:10.1002/qj.1978, http://dx.doi.org/10.1002/qj.1978, 2013.

Pannekoucke, O.: Modélisation des structures locales de covariance des erreurs de prévision à l'aide des ondelettes, Ph.D. thesis, Université Paul Sabatier-Toulouse III, https://tel.archives-ouvertes.fr/tel-00285515/document, 2008.

Pannekoucke, O.: Heterogeneous Correlation Modeling Based on the Wavelet Diagonal Assumption and on the Diffusion Operator, Monthly Weather Review, 137, 2995–3012, doi:10.1175/2009MWR2783.1, 2009.

Pannekoucke, O. and Massart, S.: Estimation of the local diffusion tensor and normalization for heterogeneous correlation modelling using a diffusion equation, Quarterly Journal of the Royal Meteorological Society, 134, 1425–1438, doi:10.1002/qj.288, http://dx.doi.org/10.1002/qj.288, 2008.

Pannekoucke, O., Emili, E., and Thual, O.: Modelling of local length-scale dynamics and isotropizing deformations, Quarterly Journal of the Royal Meteorological Society, 140, 1387–1398, doi:10.1002/qj.2204, 2014.

Pannekoucke, O., Berre, L., and Desroziers, G.: Filtering Properties of Wavelets For Local Background-Error Correlations, Quarterly Journal of the Royal Meteorological Society, 133, 363–379, doi:10.1002/qj.33, 2007.

Parrish, D. F. and Derber, J. C.: The National Meteorological Center's Spectral Statistical-Interpolation Analysis System, Monthly Weather Review, 120, 1747–1763, doi:10.1175/1520-0493(1992)120<1747:TNMCSS>2.0.CO;2, 1992.

Purser, R. J., Wu, W.-S., Parrish, D. F., and Roberts, N. M.: Numerical Aspects of the Application of Recursive Filters to Variational Statistical Analysis. Part II: Spatially Inhomogeneous and Anisotropic General Covariances, Monthly Weather Review, 131, 1536–1548, doi:10.1175//2543.1, http://dx.doi.org/10.1175//2543.1, 2003.

Raynaud, L. and Pannekoucke, O.: Sampling properties and spatial filtering of ensemble background-error length-scales, Quarterly Journal of the Royal Meteorological Society, 139, 784–794, doi:10.1002/qj.1999, 2013.

Raynaud, L., Berre, L., and Desroziers, G.: Objective filtering of ensemble-based background-error variances, Quarterly Journal of the Royal Meteorological Society, 135, 1177–1199, doi:10.1002/qj.438, 2009.

Varella, H., Berre, L., and Desroziers, G.: Diagnostic and impact studies of a wavelet formulation of background-error correlations in a global model, Quarterly Journal of the Royal Meteorological Society, 137, 1369–1379, doi:10.1002/qj.845, 2011.

Weaver, A. and Courtier, P.: Correlation modelling on the sphere using a generalized diffusion equation, Quarterly Journal of the Royal Meteorological Society, 127, 1815–1846, doi:10.1002/qj.49712757518, 2001.

Weaver, A. T., Deltel, C., Machu, E., Ricci, S., and Daget, N.: A multivariate balance operator for variational ocean data assimilation, Quarterly Journal of the Royal Meteorological Society, 131, 3605–3625, doi:10.1256/qj.05.119, 2005.